Journal of Computational and Applied Mathematics 256 (2014) 114–120

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Recovering low-rank matrices from corrupted observations via the linear conjugate gradient algorithm

Zheng-Fen Jin [a,b], Qiuyu Wang [c,*], Zhongping Wan [a]

[a] School of Mathematics and Statistics, Wuhan University, Wuhan 430072, PR China

[b] Department of Mathematics, Henan University, Kaifeng 475001, PR China

[c] College of Software, Henan University, Kaifeng 475001, PR China

ARTICLE INFO

Article history:
Received 24 August 2012
Received in revised form 13 April 2013

Keywords:
Nuclear norm minimization
Conjugate gradient method
Alternating direction method
Singular value thresholding
Augmented Lagrangian function

ABSTRACT

The matrix nuclear norm minimization problem has received much attention in recent years, largely because its highly related to the matrix rank minimization problem arising from controller design, signal processing and model reduction. The alternating direction method is a very popular way to solve this problem due to its simplicity, low storage, practical computation efficiency and nice convergence properties. In this paper, we propose an alternating direction method, where one variable is determined explicitly, and the other variable is computed by a linear conjugate gradient algorithm. At each iteration, the method involves a singular value thresholding and its convergence result is guaranteed in this literature. Extensive experiments illustrate that the proposed algorithm compares favorable with the state-of-the-art algorithms FPCA and IADM_BB which were specifically designed in recent years.

1. Introduction

Let $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^p$ be a linear map and $b \in \mathbb{R}^p$. The linear constrained nuclear norm minimization problem is to find $X \in \mathbb{R}^{m\times n}$ such that

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_*, \quad \text{s.t. } \mathcal{A}(X) = b. \tag{1}$$

Assuming that $\text{rank}(X) = r$, the nuclear norm of X is defined as $\|X\|_* = \sum_{i=1}^{r} \sigma_i(X)$, where $\sigma_i(X)$ are the singular values of X. Problem (1) is a commonly-used convex relation of the matrix rank minimization problem which arises from various areas such as machine learning, statistics, engineering and so on [1]. In practical applications, if some noises are contaminated, problem (1) should be relaxed to the inequality constrained problem

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_*, \quad \text{s.t. } \|\mathcal{A}(X) - b\|_2 \le \delta, \tag{2}$$

where $\delta \ge 0$ is the noise level. Model (2) is equal to (1) as $\delta = 0$. Another variant of the model (2) is the nuclear norm regularized least square

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_* + \frac{\gamma}{2}\|\mathcal{A}(X) - b\|_2^2, \tag{3}$$

* Corresponding author. Tel.: +86 378 3881696; fax: +86 378 3881696.
E-mail address: wqy@henu.edu.cn (Q. Wang).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.07.009

where $\gamma > 0$ is to balance the two terms for minimization. From the optimization theory, with proper choices on δ and γ, it is known that the solutions of (2) and (3) are accordant, and both problems are equivalent in some sense.

Many algorithms have been developed in recent years to solve matrix nuclear norm problems with different types. Cai, Candès, and Shen [2] presented a singular value thresholding algorithm to solve an approximate version of (1). Ma, Goldfarb, and Chen [3] extended the well-known fixed point continuation algorithm [4] to solve the penalty function of (1) with approximate singular value decomposition (FPCA). Yang and Yuan [5] presented an alternating direction method (ADM) to solve the nuclear norm regularized least square (3) as well as the models (1) and (2). In their algorithm, a linearized technique is used, and both subproblems admit explicit solutions. Instead of the linearized technique, the ADM algorithm of Xiao and Jin [6] solves one of the subproblems iteratively by the Barzilai–Borwein gradient method [7].

In this paper, we further consider the ADM algorithm and use it to solve matrix nuclear norm minimization problems. Our algorithm is similar to the ADM algorithm of Xiao and Jin; that is, to reformulate (1) as an equivalent model with an auxiliary variable and to minimize the corresponding augmented Lagrangian function alternatively. One variable is computed by using a singular value shrinkage, while the other variable is computed by solving a linear system with a conjugate gradient method [8]. Although the linear conjugate gradient method is classic and its convergence properties have been well studied, its remarkable effectiveness in matrix nuclear norm minimization is verified in our work.

We organize the rest of the paper as follows. In Section 2, we briefly review the classic ADM method and construct our method subsequently. Then in Section 3, we present some numerical results to illustrate the efficiency of the proposed algorithm. Finally, we give some conclusions in Section 4.

2. Algorithm

The earliest ADM results were from Glowinski and Marrocco [9], and Gabay and Mercier [10]. ADM is designed to solve the following separable convex optimization problem

$$\min_{x,y} \ \theta_1(x) + \theta_2(y) \quad \text{s.t. } Ax + By = c, \tag{4}$$

where $\theta_1 : \mathbb{R}^s \to \mathbb{R}$, $\theta_2 : \mathbb{R}^t \to \mathbb{R}$ are convex functions, and $A \in \mathbb{R}^{l\times s}$, $B \in \mathbb{R}^{l\times t}$, and $c \in \mathbb{R}^l$. The corresponding augmented Lagrangian function is

$$\mathcal{L}_{\mathcal{A}}(x, y, \lambda) = \theta_1(x) + \theta_2(y) - \lambda^\top(Ax + By - c) + \frac{\beta}{2}\|Ax + By - c\|_2^2, \tag{5}$$

where $\lambda \in \mathbb{R}^l$ is the Lagrangian multiplier and $\beta > 0$ is a penalty parameter. The classical ADM method is to minimize (5) first with respect to x, then with respect to y, and then update λ subsequently, i.e.,

$$\begin{cases} x_{k+1} \leftarrow \arg\min_x \mathcal{L}_{\mathcal{A}}(x, y_k, \lambda_k), \\ y_{k+1} \leftarrow \arg\min_y \mathcal{L}_{\mathcal{A}}(x_{k+1}, y, \lambda_k), \\ \lambda_{k+1} \leftarrow \lambda_k - \beta[Ax_{k+1} + By_{k+1} - c]. \end{cases}$$

The main advantage of ADM is to make full use of the separability structure of the objection function $\theta_1(x) + \theta_2(y)$. For a theoretical analysis of ADM, we can refer to Proposition 4.2 and its proof in book [11, Chapter 3, p. 256]. Recent developments in ADM can be found in [12–21].

Based on the above analysis, we now re-consider problem (1). By introducing an auxiliary variable Y, the original model (1) is equivalently transformed into

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_*, \quad \text{s.t. } X - Y = 0, \qquad \mathcal{A}(Y) = b. \tag{6}$$

Its augmented Lagrangian function is

$$\mathcal{L}_{\mathcal{A}}(X, Y, Z, z) = \|X\|_* - \langle Z, X - Y\rangle + \frac{\mu}{2}\|X - Y\|_F^2 - \langle z, \mathcal{A}(Y) - b\rangle + \frac{\gamma}{2}\|\mathcal{A}(Y) - b\|_2^2, \tag{7}$$

where $z \in \mathbb{R}^p$ and $Z \in \mathbb{R}^{m\times n}$ are multipliers of equality constraints, $\mu > 0$ and $\gamma > 0$ are penalty parameters, and $\langle\cdot\rangle$ denotes the standard trace inner product for the matrix, or the standard inner product for vectors. For fixed (X_k, Y_k), the next pair (X_{k+1}, Y_{k+1}) is determined by

$$X_{k+1} \leftarrow \arg\min_X \mathcal{L}_{\mathcal{A}}(X, Y_k, Z_k, z_k), \tag{8}$$

$$Y_{k+1} \leftarrow \arg\min_Y \mathcal{L}_{\mathcal{A}}(X_{k+1}, Y, Z_k, z_k), \tag{9}$$

$$Z_{k+1} \leftarrow Z_k - \mu[X_{k+1} - Y_{k+1}], \tag{10}$$

$$z_{k+1} \leftarrow z_k - \gamma[\mathcal{A}(Y_{k+1}) - b]. \tag{11}$$